*Indian J. Genet*., 81(1): 119-126 (2021)
DOI: 10.31742/IJGPB.81.1.13

# New hyper-variable SSRs for diversity analysis in mango (*Mangifera indica* L.)

**Manish Srivastav*, Sanjay K. Singh, Jai Prakash, Rakesh Singh[1], Neha Sharma, Sridhar Ramchandra, Rajni Devi, Ankit Gupta, Ajay K. Mahto[2], Pawan K. Jayaswal[2], Sangeeta Singh[2] and Nagendra K. Singh[2]**

Division of Fruits and Horticultural Technology, ICAR-Indian Agricultural Research Institute, New Delhi 110 012; [1]Division of Genomic Resources, ICAR-National Bureau of Plant Genetic Resources, New Delhi 110 012; [2]ICAR-National Institute for Plant Biotechnology, New Delhi 110 012

(Received: October 2020; Revised: January 2021; Accepted: February 2021)

**Abstract**

**Whole genome sequence in mango offers unprecedented opportunities for genomics assisted crop improvement *via* enabling access to genome-wide genetic markers. In the present study, simple sequence repeats (SSRs) were predicted from genome sequence of mango. Based on the SSR length ($\leq$ 50 bp), highly-variable mango SSRs (MSSRs) were sorted. A sub-set of 129 MSSRs was validated on a set of 24 diverse mango genotypes yielding marker validation and polymorphism of 88.96 and 85.27 per cent, respectively. One hundred and ten polymorphic markers were identified for the present set of mango genotypes. Polymorphic information content (PIC) ranged from 0.10 to 0.78 and the highest value was observed with MSSR133. The mean PIC value was 0.40 but 33 MSSR markers showed PIC values $\geq$ 0.5, suggesting that these markers can efficiently measure genetic diversity and serve for mapping of quantitative trait loci (QTLs) in mango. MSSRs data was further used for diversity analysis of mango genotypes belonging to different agro-ecological conditions including chance seedlings, landraces, exotic and indigenous germplasm and hybrids. Cluster analysis using UPGMA of 24 mango genotypes revealed that these MSSRs were informative in diversity analysis and distinguished mango genotypes based on their origin, parentage and embryo types. A novel set of 110 hyper-variable SSR markers have been added to the mango repository depicting usefulness in discriminating closely related mango germplasm and their use in mango improvement programme.**

**Key words:** *Mangifera indica*, hyper-variable SSR, validation, diversity analysis

## Introduction

Mango (*Mangifera indica* L.) originated in the Indo-Myanmar region (Yonemori et al. 2002) and gradually spread to the tropical and sub-tropical regions of the world. Presently, India represents the biggest collection of *M. indica* germplasm in the world. Mango being considered 'King of fruits' found to be economically important due to its excellent qualities (attractive colour, sweet taste and mouth watering flavour) and nutritional composition (vitamins, minerals, edible fibre, and phytochemicals) (Kim et al. 2009). In India, in terms of area, mango ranks first and third in respect of production and produces about 18.43 m t of mangoes annually from an area of 2.52 m ha. India is the largest producer and the second largest exporter of this tropical fruit. India's share in the world's mango production is around 56 per cent. During 2013-14, India exported 41,280 mt of fresh mango valued at 2,854.3 million Rupees, whereas in 2018-19 it exported 46510.27 m t fresh mangoes accounting for Rs. 406.45 crores (60.26 million US$).

Majority of mango germplasm possess desirable morphological and horticultural traits, however, there is a confusion and uncertainty concerning the identity of the mango genotypes and their exact parentage (Krishna and Singh 2007). Therefore, before initiating any breeding programme, it is essential for a breeder to accurately estimate the relationship between different genotypes and to select the most appropriate parents. Further, molecular characterization, due to their reliability and repeatability is considered most advantageous and facilitate an efficient management and utilization of the available germplasm.

Molecular markers are among the pre-requisites

*Corresponding author's e-mail: msrivastav@iari.res.in
Published by the Indian Society of Genetics & Plant Breeding, A-Block, F2, First Floor, NASC Complex, IARI P.O., Pusa Campus, New Delhi 110 012; Online management by www.isgpb.org; indianjournals.com

to accelerate the breeding program through genomics assisted breeding. DNA markers previously used for characterization of genetic diversity in mango include random amplified polymorphic DNA (RAPDs) (Schnell et al. 1995; Lopez-Valenzuela et al. 1997; Ravishankar et al. 2000; Kumar et al. 2001; Karihaloo et al. 2003; Bajpai et al. 2008; Roy et al. 2011); amplified fragments length polymorphism (AFLP) (Eiadthong et al. 2000; Kashkush et al. 2001); inter-simple sequence repeats (Eiadthong et al.1999; Pandit et al. 2007) and simple sequence repeats (Duval et al. 2005; Viruel et al. 2005; Schnell et al. 2006; Ravishankar et al. 2011; Dillon et al. 2013). Among all the marker types, SSRs have always become the markers of choice due to their co-dominant nature, polymorphic behaviour and higher reproducibility (Rajwant et al. 2011; Kumari et al. 2020).

In the recent past, significant progress has been made in the area of whole genome sequencing of mango (Singh et al. 2014; Singh et al. 2016; Luo et al. 2016; Kuhn et al. 2017). These genomic resources in mango offers unprecedented opportunities for genomics assisted crop improvement *via* enabling access to genome-wide genetic markers. SSRs identified from whole genome sequences in mango have been divided into two classes based on their repeat length and potential as informative genetic markers, *i.e.*, class I SSRs with repeat lengths of 20 bp or higher, and class II SSRs with repeat lengths of 12–19 bp (Temnykh et al. 2001). The rationale for making the two classes was that the SSRs with larger number of repeats were more polymorphic than those with less number of repeats as reported in human (Weber 1990; Xu et al. 2000). The class I SSRs were found more polymorphic than the class II SSR and denoted as hyper-variable marker (Temnykh et al. 2001). Singh et al. (2010) also reported highest degree of polymorphism in the SSR length range of 51–70 bp, beyond which there was stabilization and then decline of polymorphism in SSRs longer than 70 bp. In the present study, our aim was to validate the new MSSRs developed from the genome sequences of mango and to test their usefulness in diversity analysis of mango germplasm.

## Materials and methods

### *Designing of simple sequence repeat markers*

RNA-seq of mango cultivars Neelum, Dashehari and their hybrid Amrapali revealed substantially higher level of heterozygosity in Amrapali over its parents and helped in developing genic simple sequence repeat (SSR) markers (Mahato et al. 2016). Hyper-variable SSRs are simple sequence repeats whose repeating unit is > 50 bp and are more polymorphic as compared to other type of SSRs (Singh et al. 2010). In present study, mono and complex type SSRs were excluded and only di, tri, tetra, penta and hexa type SSRs were used for primer designing and validation. Highly variable mango SSRs (MSSRs) and designed flanking primer using PRIMER3 were used in present investigation. A set of 24 diverse mango germplasm (Table 1) comprising of exotic, indigenous cultivars, chance seedlings, landraces and hybrids from different agro-ecological regions conserved at the field gene bank of the Division of Fruits and Horticultural Technology, ICAR- Indian Agricultural Research Institute, New Delhi were taken for validation and diversity studies.

### *DNA extraction, quantification and PCR amplification*

Newly emerged fresh leaves of different mango genotypes were collected from the Field Gene Bank of ICAR- Indian Agricultural Research Institute, New Delhi. The genomic DNA was isolated using CTAB method as described by Doyle and Doyle (1987) with minor modifications. The fresh leaf tissue was grounded in liquid nitrogen using the pre-sterilized pestle and mortar. The grounded leaf sample was then homogenized with 1.0 ml CTAB and kept in pre-sterilized centrifuge tubes. These tubes were then kept at 65$^{o}$C for 1 h and inverted at 10 min. interval. The incubated samples were then emulsified with equal volume of chloroform: isoamyl alcohol (24:1) and centrifuged at 12,000 x g for 10 min. The supernatants were transferred in the pre-sterilized centrifuge tubes; thereafter pre-chilled absolute alcohol and 3M sodium acetate was added to facilitate the precipitation of genomic DNA and incubated at 4$^{o}$C for overnight. To get the pellet of predicated DNA, the incubated tubes were centrifuged at 12,000 x g for 10 min. and the DNA pellet was washed with 70% chilled ethanol twice with short spin on centrifuge. The DNA pellets were air-dried and dissolved in 50 µl of TE Buffer. To remove the RNA impurities, the diluted DNA was treated with RNase (0.5 µl) and purified. The purified DNA was quantified with 0.8% agarose gel and spectrophotometer (Nanodrop$^{TM}$, Thermo Fisher, USA). The final concentration of purified DNA was maintained at 20 ng/µl in nuclease-free water as working dilution and kept in deep freeze (–20$^{o}$C).

The PCR reaction was performed in 10 µl volume of mixture, containing 2.0 µl of 20 ng genomic DNA

**Table 1.** A list of 24 mango genotypes used for validation and diversity analysis

| S.No. | Name | Origin | Type of embryo | Progenitor/ parentage |
|---|---|---|---|---|
| 1. | Ratna | West India | Mono | Neelum x Alphanso |
| 2. | Rataul | North India | Mono | Chance seedling |
| 3. | Khasulkhas | North India | Mono | Chance seedling |
| 4. | Langra | North India | Mono | Chance seedling |
| 5. | Lucknow Safeda | North India | Mono | Chance seedling |
| 6. | Zill | Florida, USA | Mono | Haden |
| 7. | Mallika | North India | Mono | Neelum x Dashehari |
| 8. | Totapari | South India | Mono | Chance seedling |
| 9. | Suvernarekha | South India | Mono | Chance seedling |
| 10. | Iturba | Mexico | Mono | Chance seedling |
| 11 | Kurukkan | South India | Poly | Chance seedling |
| 12. | Olour | South India | Poly | Chance seedling |
| 13. | Neelum | South India | Mono | Chance seedling |
| 14 | Alphanso | West India | Mono | Chance seedling |
| 15. | Kalepad | South India | Mono | Chance seedling |
| 16. | Amrapali | North India | Mono | Dashehari x Neelum |
| 17. | Chausa | North India | Mono | Chance seedling |
| 18. | Dashehari | North India | Mono | Chance seedling |
| 19. | Tommy Atkins | Florida, USA | Mono | Haden |
| 20. | Fernandin | West India | Mono | Chance seedling |
| 21. | Maya | Israel | Mono | Open pollinated seedling |
| 22. | Kesar | West India | Mono | Chance seedling |
| 23. | Manjeera | South India | Mono | Rumani x Neelum |
| 24. | Sindhu | West India | Mono | Ratna x Alphanso |

and 5 µl Ready PCR Mix (One PCR$^{TM}$, GeneDireX), 1.0 µl (0.5 µl forward + 0.5 µl reverse) of 100 pmol primer and 2.0 µl of molecular grade water to raise the final volume of 10 µl. The PCR amplification was carried out in thermalcycler (Bio Rad, USA). Thermal profiling was setup with initial denaturation temperature of 94 °C for 3 min. followed by the 35 cycles of denaturation (94 °C for 60s), annealing (50-55 °C for 60s) and extension (72 °C for 60s) and final extension at 72 °C for 10 min. The amplified SSR fragments were size separated on 4% agarose gel (Lonza, Switzerland) containing ethidium bromide (2.0 ìg/ 100 ml) in 1X TAE buffer and photographed on gel documentation system (BioSystematica, UK).

### *SSR allele scoring and data analysis*

The total number of monomorphic and polymorphic bands which were clear, unambiguous and reproducible was scored for all SSRs. Data scoring was carried out by using a binary number system for the presence or absence of each fragment in each sample. Measure of degree of similarity among 24 mango genotypes was established as a percentage of polymorphic bands, and a matrix of genetic similarity compiled with the NTSYS programme SIMINT using the product-moment correlation coefficient (Rohlf 2000). Cluster analysis was performed with the NTSYS programme SHAN using unweighted pair-group arithmetic average (UPGMA) and a dendrogram representing relationship among 24 mango genotypes was generated. The genetic diversity indices, *viz*., major allelic frequency, gene diversity or expected heterozygosity (Nei 1972), observed heterozygosity and polymorphism information content of each SSR locus were calculated using Power Marker ver. 3.25 (Liu and Muse 2004).

## Results

### *Validation of hyper-variable mango SSRs*

In the present study, 145 MSSRs have been used for their validation in a set of 24 diverse mango genotypes. A total of 2,976 fragments and 210 alleles in size from 100 to 320 bp were amplified by 129 MSSR primers, while remaining 16 MSSRs did not amplify in the given set of mango germplasm. PCR conditions were optimized by testing various annealing temperatures between 48.0° to 60.0°C for these failed primers even then these primers did not amplify. Out of 129 validated primers, 110 (85.27%) showed polymorphism, while 19 (14.73%) were found to be monomorphic. Number of alleles ranged from 1 to 4. The polymorphism information content (PIC) ranged from 0.10 to 0.78 and the highest PIC was observed with the SSR133 (Table 2; Fig. 1a-c) with mean value of 0.40. The SSR loci (33), *viz*., MSSR4, MSSR21, MSSR27, MSSR29, MSSR31, MSSR37, MSSR38, MSSR41, MSSR43, MSSR44, MSSR45, MSSR48, MSSR52, MSSR53, MSSR56, MSSR60, MSSR76, MSSR79, MSSR91, MSSR92, MSSR94, MSSR95, MSSR104, MSSR105, MSSR107, MSSR111, MSSR112, MSSR114, MSSR117, MSSR128, MSSR133, MSSR141 and MSSR144 had PIC values $\geq$ 0.50, and another 17 MSSRs showed PIC values between 0.4 to 0.49 (Table 2). Thus, representing high discriminating power of the selected SSRs and can efficiently measure genetic diversity and serve for mapping of quantitative trait loci (QTLs) in mango.

### *Descriptive diversity statistics of SSR loci*

The scores of 110 SSR loci were used to determine the genetic diversity analysis among the mango genotypes including the indigenous, exotic germplasm from different ecological conditions. The diversity indices, *viz*., major allelic frequency, expected heterozygosity, observed heterozygosity and polymorphic information content were calculated for each SSR locus amplified for the studied mango genotypes. Allelic size range varied from 100 to 320 bp among the 110 SSR loci. The major allelic frequency among the 110 SSR loci ranged from 0.29 (MSSR133) to 0.95 (MSSR132) with an average of 0.65. The maximum expected heterozygosity (0.78) was noted for SSR locus MSSR133, while minimum (0.10) for SSR locus MSSR132 with mean value of 0.45 among the 110 SSRs. The observed heterozygosity ranged between 0.04 with SSR loci MSSR109 to 0.88 with MSSR117 with an average of 0.35 with 110 SSR loci. The average PIC value was 0.40 among the 110 SSR loci and the highest PIC (0.78) was recorded for locus MSSR133, while minimum (0.10) for MSSR132 (Table 2).

### *Diversity analysis of mango genotypes*

The 110 SSR data was further used for similarity analysis and tree generation. UPGMA analysis of 110 MSSR data revealed that 24 mango genotypes grouped into two main clusters. The genetic distance for the genotypes ranged from 0.66 to 0.91. Mango varieties from Western India such as Ratna, Manjeera, Sindhu, Alphanso and Kesar grouped together in Cluster I. Ratna is a cross of Neelum and Alphanso, while Sindhu is a hybrid developed by backcrossing Ratna with Alphanso. Ratna showed more than 75% similarity with Alphanso and Sindhu using 110 SSR loci data. Cluster II represented the remaining 19 genotypes was further divided in two sub-groups.

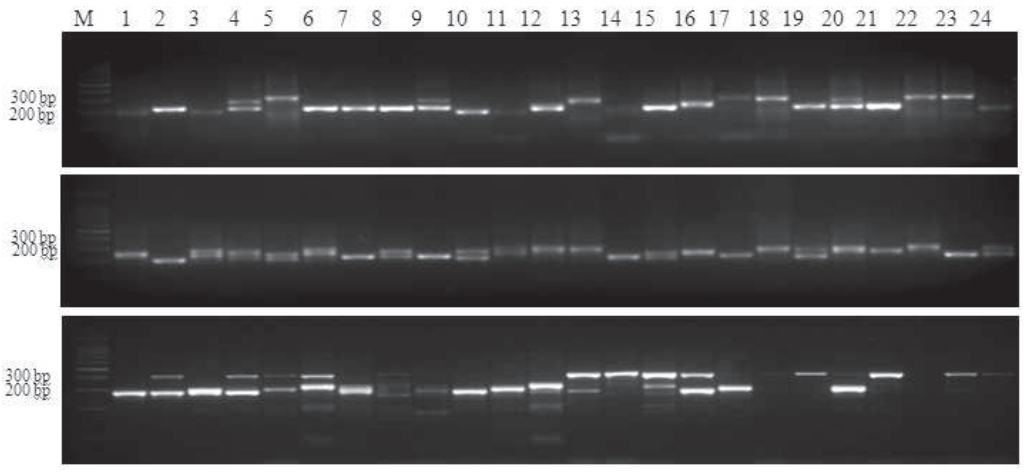


**Fig. 1. Mango SSR profiles of 24 diverse mango germplasm. , A. MSSR53, B. MSSR58 and C. MSSR133. M- Marker (100 bp), Lane 1-24: 1-Ratna, 2-Rataul, 3-Khasulkhas, 4-Langra, 5-Lucknow Safeda, 6-Zill, 7-Mallika, 8-Totapari, 9-Suvernarekha, 10-Iturbe, 11-Kurukkan, 12-Olour, 13-Neelum, 14-Alphanso, 15 Kalepad, 16-Amrapali, 17-Chausa, 18-Dashehari, 19-Tommy Atkins 20-Fernandin, 21-Maya, 22-Kesar, 23-Manjeera and 24-Sindhu**

**Table 2.** Details of polymorphic MSSR loci used for genotyping of 24 mango genotypes along with their major allelic frequencies (Maf), expected hetrozygocity (He), observed heterozygosity and (Ho) and polymorphism information content (PIC)

| S. No. | SSR | $M_{af}$ | He | Ho | PIC | S. No. | SSR | $M_{af}$ | He | Ho | PIC |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. | SSR1 | 0.46 | 0.58 | 0.08 | 0.42 | 56. | SSR77 | 0.57 | 0.49 | 0.69 | 0.37 |
| 2. | SSR2 | 0.84 | 0.22 | 0.18 | 0.21 | 57. | SSR79 | 0.54 | 0.50 | 0.83 | 0.53 |
| 3. | SSR3 | 0.65 | 0.51 | 0.21 | 0.45 | 58. | SSR83 | 0.90 | 0.19 | 0.21 | 0.37 |
| 4. | SSR4 | 0.50 | 0.62 | 0.38 | 0.53 | 59. | SSR84 | 0.90 | 0.19 | 0.21 | 0.30 |
| 5. | SSR5 | 0.94 | 0.12 | 0.13 | 0.20 | 60. | SSR88 | 0.56 | 0.49 | 0.81 | 0.37 |
| 6. | SSR6 | 0.90 | 0.19 | 0.13 | 0.22 | 61. | SSR89 | 0.70 | 0.42 | 0.30 | 0.33 |
| 7. | SSR7 | 0.91 | 0.13 | 0.14 | 0.15 | 62. | SSR90 | 0.52 | 0.50 | 0.76 | 0.37 |
| 8. | SSR8 | 0.62 | 0.47 | 0.85 | 0.37 | 63. | SSR91 | 0.40 | 0.69 | 0.54 | 0.58 |
| 9. | SSR10 | 0.57 | 0.59 | 0.09 | 0.48 | 64. | SSR92 | 0.48 | 0.67 | 0.33 | 0.68 |
| 10. | SSR11 | 0.58 | 0.44 | 0.81 | 0.37 | 65. | SSR93 | 0.58 | 0.57 | 0.08 | 0.44 |
| 11. | SSR12 | 0.88 | 0.15 | 0.16 | 0.20 | 66. | SSR94 | 0.73 | 0.40 | 0.54 | 0.50 |
| 12. | SSR13 | 0.74 | 0.33 | 0.24 | 0.22 | 67. | SSR95 | 0.75 | 0.38 | 0.50 | 0.58 |
| 13. | SSR14 | 0.44 | 0.63 | 0.41 | 0.31 | 68. | SSR96 | 0.71 | 0.41 | 0.58 | 0.42 |
| 14. | SSR18 | 0.79 | 0.29 | 0.41 | 0.25 | 69. | SSR97 | 0.83 | 0.29 | 0.15 | 0.25 |
| 15. | SSR19 | 0.61 | 0.48 | 0.78 | 0.36 | 70. | SSR98 | 0.57 | 0.49 | 0.41 | 0.37 |
| 16. | SSR21 | 0.52 | 0.60 | 0.54 | 0.52 | 71. | SSR99 | 0.85 | 0.26 | 0.30 | 0.22 |
| 17. | SSR23 | 0.69 | 0.47 | 0.21 | 0.40 | 72. | SSR100 | 0.83 | 0.27 | 0.33 | 0.23 |
| 18. | SSR25 | 0.61 | 0.47 | 0.75 | 0.36 | 73. | SSR101 | 0.84 | 0.25 | 0.21 | 0.18 |
| 19. | SSR26 | 0.56 | 0.49 | 0.87 | 0.37 | 74. | SSR103 | 0.83 | 0.28 | 0.33 | 0.24 |
| 20. | SSR27 | 0.40 | 0.66 | 0.38 | 0.58 | 75. | SSR104 | 0.65 | 0.49 | 0.15 | 0.50 |
| 21. | SSR29 | 0.58 | 0.55 | 0.21 | 0.60 | 76. | SSR105 | 0.55 | 0.53 | 0.68 | 0.61 |
| 22. | SSR30 | 0.63 | 0.53 | 0.21 | 0.49 | 77. | SSR107 | 0.63 | 0.54 | 0.13 | 0.51 |
| 23. | SSR31 | 0.46 | 0.64 | 0.25 | 0.50 | 78. | SSR108 | 0.86 | 0.25 | 0.18 | 0.22 |
| 24. | SSR33 | 0.81 | 0.30 | 0.38 | 0.44 | 79. | SSR109 | 0.58 | 0.57 | 0.04 | 0.41 |
| 25. | SSR34 | 0.73 | 0.41 | 0.13 | 0.47 | 80. | SSR110 | 0.46 | 0.62 | 0.33 | 0.38 |
| 26. | SSR37 | 0.46 | 0.60 | 0.08 | 0.55 | 81. | SSR111 | 0.83 | 0.29 | 0.08 | 0.58 |
| 27. | SSR38 | 0.42 | 0.63 | 0.00 | 0.51 | 82. | SSR112 | 0.77 | 0.41 | 0.08 | 0.50 |
| 28. | SSR39 | 0.50 | 0.59 | 0.00 | 0.31 | 83. | SSR113 | 0.82 | 0.30 | 0.12 | 0.25 |
| 29. | SSR40 | 0.58 | 0.49 | 0.84 | 0.37 | 84. | SSR114 | 0.48 | 0.65 | 0.71 | 0.72 |
| 30. | SSR41 | 0.42 | 0.58 | 0.21 | 0.50 | 85. | SSR115 | 0.72 | 0.40 | 0.13 | 0.32 |
| 31. | SSR42 | 0.73 | 0.41 | 0.29 | 0.38 | 86. | SSR116 | 0.46 | 0.57 | 0.00 | 0.30 |
| 32. | SSR43 | 0.54 | 0.56 | 0.75 | 0.50 | 87. | SSR117 | 0.56 | 0.49 | 0.88 | 0.53 |
| 33. | SSR44 | 0.54 | 0.58 | 0.67 | 0.52 | 88. | SSR118 | 0.46 | 0.64 | 0.17 | 0.44 |
| 34. | SSR45 | 0.60 | 0.48 | 0.79 | 0.55 | 89. | SSR119 | 0.61 | 0.48 | 0.27 | 0.36 |
| 35. | SSR46 | 0.67 | 0.44 | 0.25 | 0.40 | 90. | SSR121 | 0.81 | 0.30 | 0.12 | 0.26 |
| 36. | SSR47 | 0.71 | 0.41 | 0.17 | 0.48 | 91. | SSR122 | 0.72 | 0.41 | 0.12 | 0.22 |
| 37. | SSR48 | 0.69 | 0.43 | 0.21 | 0.51 | 92. | SSR123 | 0.58 | 0.49 | 0.68 | 0.37 |
| 38. | SSR49 | 0.79 | 0.33 | 0.42 | 0.28 | 93. | SSR124 | 0.82 | 0.30 | 0.19 | 0.25 |
| 39. | SSR52 | 0.38 | 0.66 | 0.17 | 0.51 | 94. | SSR126 | 0.53 | 0.50 | 0.24 | 0.37 |
| 40. | SSR53 | 0.69 | 0.45 | 0.54 | 0.50 | 95. | SSR127 | 0.63 | 0.50 | 0.58 | 0.48 |
| 41. | SSR55 | 0.63 | 0.53 | 0.17 | 0.38 | 96. | SSR128 | 0.81 | 0.32 | 0.21 | 0.37 |
| 42. | SSR56 | 0.54 | 0.56 | 0.75 | 0.71 | 97. | SSR130 | 0.75 | 0.40 | 0.08 | 0.50 |
| 43. | SSR57 | 0.57 | 0.49 | 0.32 | 0.37 | 98. | SSR131 | 0.56 | 0.58 | 0.29 | 0.38 |
| 44. | SSR58 | 0.58 | 0.52 | 0.08 | 0.45 | 99. | SSR132 | 0.95 | 0.10 | 0.11 | 0.10 |
| 45. | SSR60 | 0.63 | 0.52 | 0.17 | 0.52 | 100. | SSR133 | 0.29 | 0.78 | 0.63 | 0.78 |
| 46. | SSR61 | 0.60 | 0.48 | 0.36 | 0.37 | 101. | SSR134 | 0.53 | 0.50 | 0.33 | 0.37 |
| 47. | SSR62 | 0.68 | 0.44 | 0.33 | 0.34 | 102. | SSR135 | 0.73 | 0.42 | 0.13 | 0.22 |
| 48. | SSR63 | 0.67 | 0.44 | 0.54 | 0.35 | 103. | SSR136 | 0.71 | 0.41 | 0.36 | 0.33 |
| 49. | SSR65 | 0.60 | 0.48 | 0.80 | 0.37 | 104. | SSR137 | 0.58 | 0.49 | 0.15 | 0.37 |
| 50. | SSR67 | 0.75 | 0.39 | 0.08 | 0.23 | 105. | SSR138 | 0.93 | 0.12 | 0.13 | 0.12 |
| 51. | SSR68 | 0.71 | 0.41 | 0.49 | 0.33 | 106. | SSR139 | 0.69 | 0.43 | 0.60 | 0.34 |
| 52. | SSR69 | 0.90 | 0.19 | 0.21 | 0.30 | 107. | SSR141 | 0.56 | 0.54 | 0.71 | 0.60 |
| 53. | SSR73 | 0.58 | 0.56 | 0.08 | 0.48 | 108. | SSR142 | 0.88 | 0.21 | 0.23 | 0.18 |
| 54. | SSR74 | 0.52 | 0.57 | 0.21 | 0.44 | 109. | SSR143 | 0.68 | 0.40 | 0.41 | 0.31 |
| 55. | SSR76 | 0.50 | 0.61 | 0.67 | 0.60 | 110. | SSR144 | 0.48 | 0.63 | 0.04 | 0.50 |

Amrapali and Mallika hybrids grouped together with their parents Neelum and Dashehari in sub-group $II_a$. It was also evident that Mallika had more similarity with parent Neelum than Dashehari. However, Amrapali showed more similarity with its maternal parent Dashehari than Neelum.

Cluster $II_b$ comprising of 15 genotypes was further divided in two more sub-groups. In first sub-group of $II_b$ cluster mango genotypes, *viz*., Rataul, Zill, Khasulkhas and Lucknow Safeda showed similarity than other genotypes. In second sub-group of $II_b$ cluster, 11 genotypes including Langra, Fernandin, Kurukkan, Olour, Kalepad, Tommy Atkins, Maya, Totapari, Suvernarekha and Iturba grouped together. The maximum similarity (0.91) was noted between exotic mango genotypes Tommy Atkins from USA and Israeli Maya. Polyembryonic genotypes, *viz*., Kurukkan and Olour showed similarity (0.84) and north Indian Langra had similarity with Fernandin from Goa region (0.83). It was evident that the present set of 110 MSSRs were capable of differentiating the mango genotypes on the basis of their origin and embryo type and thus found to be useful for diversity analysis of mango germplasm.

## Discussion

Whole genome sequencing of mango cv. Amrapali has offered access to genome wide SSRs for their diverse applications in genomics and breeding. Here, we describe a genome wide set of 129 validated highly-variable MSSR markers with repeat lengths of 51–70 bp for their consistent amplification and high polymorphism. A set of selected 145 hyper-variable SSR markers were validated with success rate of 88.96%, which was comparable to earlier reported by Shareefa (2008) and Nayak (2010). In a similar way, high level of MSSR polymorphism has been observed, wherein out of 210 bands generated 174 were polymorphic. The per cent polymorphism reported here was also greater than reported earlier in mango. For instance, the percent polymorphism shown by SSR markers was reported to be 54% (Gitahi et al. 2016), 74.1% (Surapaneni et al. 2012) and 71.02% (Shareefa 2008). A greater polymorphism percentage coupled with the consistent amplification patterns render these MSSR markers to be highly suitable for mango genotyping.

Descriptive diversity statistics of SSR loci indicated that the observed heterozygosity ranged from 0.04 (MSSR109) to 0.88 (MSSR117), which clearly signifying the potential of hyper-variable MSSR markers in deciphering existing allelic diversity. Allelic size ranged from 100 to 320 bp among 110 SSR loci. Earlier studies have suggested that the rate of expansion mutations is constant for all the loci but the rate of contraction mutations increases exponentially with the SSR repeat length (Xu et al. 2000). In this investigation, 110 SSR loci covering the whole genome of mango were utilized for diversity analysis. The diversity indices were congruent with the results obtained by Shareefa et al. (2008) and Nayak et al. (2010). Similarly, the average PIC value of the current analysis corroborates with previously reported mean PIC values for SSR markers in mango. High PIC value is directly proportional to the potential of revealing allelic variation. The average PIC value of SSR markers tested by different researchers vary with number of SSR markers used and also number of genotypes tested.

Results from the present study clearly indicated vast genetic diversity among the selected mango genotypes. Mango genotypes from Western India grouped together; while Amrapali and Mallika hybrids grouped together with their parents and polyembryonic genotypes showed significant similarity among themselves. It was apparent that the present set of MSSRs was able to differentiate the mango genotypes, and found to be highly useful in diversity analysis of mango germplasm. The clustering patterns resulting from model and distance based approaches will guide mango breeders for selection of the most diverse parental lines in future breeding programmes. The relevance of hyper-variable SSR markers to plant breeding is well described in various field crops (Singh et al. 2010; Narshimulu et al. 2011). It is to be noted that a novel set of 110 validated hyper-variable MSSR markers have been added which implicates them in the assessment of genetic variation of diverse mango genotypes. A greater polymorphism percentage coupled with the consistent amplification patterns renders these SSR markers highly suitable for mango genotyping using simple laboratory equipment and help mango researchers for diverse molecular applications including marker-assisted selection.

## Authors' contribution

Conceptualization of research (MS, GKS, NKS); Designing of the experiments (MS, SKS, NKS); Contribution of experimental materials (MS, SKS, JP, AKM, NKS); Execution of field/lab experiments and data collection (NS, SR, RD, AG, MS, SKS); Analysis

of data and interpretation (NS, SS, PKJ, RS, MS, SKS); Preparation of manuscript (MS, SKS, SS, PKJ, AKM).

**Declaration**

The authors declare no conflict of interest.

**Acknowledgements**

The study was funded by Indian Council of Agricultural Research under Extramural Project titled "High density linkage map and genome sequencing of mango (*Mangifera indica* cv. Amrapali)". Authors are thankful to the National Coordinator, Network Project on Functional Genomics & Genetic Modification (Mango), Director, ICAR-NIPB, New Delhi and Director, ICAR-IARI, New Delhi for the research facilities.

**References**

Bajpai A., Srivastava N., Rajan S. and Chandra R. 2008. Genetic diversity and discrimination of mango accessions using RAPD and ISSR markers. Indian J. Hort., **65:** 377-382.

Dillon N. L., Bally I. S., Wright C. L., Hucks L., Innes D. J. and Dietzgen R. G. 2013. Genetic diversity of the Australian national mango genebank. Scientia Hortic., **150:** 213-226.

Doyle J. J. and Doyle J. L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull., **19**: 11-15.

Duval M. F., Bunel J., Sitbon C. and Risterucci A. M. 2005. Development of microsatellite markers for mango (*Mangifera indica* L.). Mol. Ecol. Notes, **5**: 824-826.

Eiadthong W., Yonemori K., Kanzaki S., Sugiura A., Utsunomiya N. and Subhadrabandhu S. 2000. Amplified fragment length poly-morphism analysis for studying genetic relationship among Mangifera species in Thailand. J. Am. Soc. Hortic. Sci., **125**: 160-164.

Eiadthong W., Yonemori K., Sugiura A., Utsunomiya N. and Subhadrabandhu S. 1999. Identification of mango cultivars of Thailand and evaluation of their genetic variation using the amplified fragments by simple sequence repeat-(SSR-) anchored primers. Sci. Hortic., **82**: 57-66.

Gitahi R., Kasili R., Kyallo M. and Kehlenbeck K. 2016. Diversity of threatened local mango landraces on smallholder farms in Eastern Kenya. Forests, Trees and Livelihoods, **25** (4): 239-254.

Karihaloo J. L., Dwivedi Y. K., Archak S. and Gaikwad A. B. 2003. Analysis of genetic diversity of Indian mango cultivars using RAPD markers. J. Hortic. Sci. Biotech. **78**: 285-289.

Kashkush K., Jinggui F., Tomer E., Hillel J. and Lavi U. 2001. Cultivar identification and genetic map of mango (*Mangifera indica*). Euphytica, **122**: 129-136.

Kim Y., Lounds-Singleton A. J. and Talcott S. T. 2009. Antioxidant phytochemical and quality changes associated with hot water immersion treatment of mangoes (*Mangifera indica* L.). Food Chem., **115**: 989-993.

Krishna H. and Singh S. K. 2007. Biotechnological advances in mango (*Mangifera indica* L.) and their future implication in crop improvement - A review. Biotechnol. Adv., **25**: 223-243.

Kuhn D. N., Bally I. S. E., Dillon N. L., Innes D., Groh A. M., Rahaman J., Ophir R., Cohen Y. and Sherman A. 2017. Genetic map of mango: A tool for mango breeding. Front. Plant Sci., **8**: 577. doi: 10.3389/fpls.2017.00577.

Kumar H. N. V., Narayanswamy P., Prasad T. D., Mukunda G. K. and Sondur S. N. 2001. Estimation of genetic diversity of commercial mango cultivars using RAPD markers. J. Hortic. Sci. Biotech., **76**: 529-533.

Kumari S., Chakraboraty S. K., Bhowmick P. K., Singh V. J. and Prasad A.S.M. 2020. Validation of hybrid rice seed vigour traits using SSR markers (*Oryza sativa* L.). Indian J. Genet., **80**(2): 204-208. DOI: 10.31742/IJGPB.80.2.11.

Liu K. and Muse S. 2004. Power Marker: new genetic data analysis software, version 2.7 (http://www.powermarker.net).

Lopez-Valenzuela J. A., Martinez O. and Paredes-Lopez O. 1997. Geographic differentiation and embryo type identification in *Mangifera indica* L. cultivars using RAPD markers. Hort. Sci., **32**: 1105-1108.

Luo C., Shu B., Yao Q., Wu H., Xu W. and Wang S. 2016. Construction of a high-density genetic map based on large-scale marker development in mango using specific-locus amplified fragment sequencing (SLAF-seq). Front. Plant Sci., **7**:1310. doi: 10.3389/fpls.2016.01310.

Mahato A. K., Sharma N., Singh A., Srivastav M., Jaiprakash, Singh S. K., Singh A. K., Sharma T. R. and Singh N. K. 2016. Leaf transcriptome sequencing for identifying genic-SSR markers and SNP heterozygosity in crossbred mango variety 'Amrapali' (*Mangifera indica* L.). PLoS ONE, **11**(10): e0164325. doi:10.1371/journal.pone.0164325.

Narshimulu G., Jamaloddin M., Vemireddy L. R., Anuradha G. and Siddiq E. 2011. Potentiality of evenly distributed hyper-variable microsatellite markers in marker-assisted breeding of rice. Plant Breed., **130**: 314-320.

Nayak D. 2010. Characterization of mango hybrids for fruit quality traits. Ph.D. thesis submitted to PG School, IARI, New Delhi.

Nei M. 1972. Genetic distance between populations. Am. Nat., **106** (949): 283-292.

Pandit S. S., Mitra S., Giri A. P., Pujari K. H., Patil B. P., Jambhale N. D. and Gupta V. S. 2007. Genetic diversity analysis of mango cultivars using inter simple sequence repeat markers. Curr. Sci., **93**:1135-1141.

Rajwant K. K., Manoj K. R., Sanjay K., Singh R. and Dhawan A. K. 2011. Microsatellite markers: an overview of the recent progress in plants. Euphytica, **177**: 309-334.

Ravishankar K. V., Anand L. and Dinesh M. R. 2000. Assessment of genetic relatedness among mango cultivars of India using RAPD markers. J. Hortic. Sci. Biotech., **75**: 198-201.

Ravishankar K. V., Mani B. H., Anand L. and Dinesh M. R. 2011. Development of new microsatellite markers from mango (*Mangifera indica*) and cross-species amplification. Am. J. Bot., **98**: e96–e99.

Rohlf F. J. 2000. NTSYS-pc: Numerical taxonomy and multivariate analysis system, version 2.2. Exeter Software, Setauket, New York, USA.

Roy S.C. and Chattopadhyay A. 2011. Evaluation of genetic diversity in mango germplasm resources using RAPD markers and characterization of cultivar Guti based on 18SrRNA gene sequence. Indian J. Genet., **71**(3): 254-261.

Schnell R. J., Brown J. S., Olano C. T., Meerow A. W., Campell R. J. and Kuhn D. N. 2006. Mango genetic diversity analysis and pedigree inferences for Florida cultivars using microsatellite markers. J. Am. Soc. Hortic., Sci. **131**: 214-224.

Schnell R. J., Ronning C. M. and Knight R. J. 1995. Identification of cultivars and validation of genetic relationships in *Mangifera indica* L. using RAPD markers. Theor. Appl. Genet., **90**: 269-274.

Shareefa M. 2008. DNA fingerprinting of mango (*Mangifera indica* L.) genotypes using molecular markers. Ph.D. thesis submitted to P.G. School, IARI, New Delhi.

Singh H., Deshmukh R. K., Singh A., Singh A. K., Gaikwad K., Sharma T. R., Mohapatra T., Singh N. K. 2010. Highly variable SSR markers suitable for rice genotyping using agarose gels. Mol. Breed., **25**: 359-364.

Singh N. K., Mahato A. K., Jayaswal P. K., Singh A., Singh S., Singh N., Rai V., Amita Mitra S.V., Gaikwad K., Sharma N., Lal S., Srivastav M., Jaiprakash, Kalindindi U., Singh S. K., Singh A. K., Khan K., Mishra R. K., Rajan S., Bajpai A., Sandhya B. S., Nischita P., Ravishankar K. V., Dinesh M. R., Kumar N., Jaiswak S., Iquebal M. A., Kumar D., Rai A., Sharma T. R. 2016. Origin, diversity and genome sequence of mango (*Mangifera indica* L.). Indian J. Hist. Sci., **51:** 355-368.

Singh N. K., Mahato A. K., Sharma N., Gaikwad K., Srivastava M., Tiwari K., Dogra V., Rawal H. C., Jayaswal P., Singh A., Rai V., Amitha Mithra S. V., Bajpai A., Dinesh M. R., Ravishankar K. V., Rajan S., Rai A., Singh A. K. and Sharma T. R. 2014. A draft genome of the King of Fruit mango (*Mangifera indica* L.). PAGXIII 2015, PAG 2014.

Surapaneni M., Vemireddy L.R., Begum H., Reddy, B.P., Neetasri C., Nagaraju J., Anwar S.Y. and Siddiq E.A. 2012. Population structure and genetic analysis of different utility types of mango (Mangifera indica L.) germplasm of Andhra Pradesh state of India using microsatellite markers. Plant Syst. Evol., doi 10.1007/s00606-013-0790-1.

Temnykh S., DeClerck G., Lukashova A., Lipovich L., Cartinhour S., McCouch S. R. 2001. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. Genome Res., **11**(8): 1441–1452.

Viruel M. A., Escribano P., Barbieri M., Ferri M. and Hormaza J. I. 2005. Fingerprinting, embryo type and geographic differentiation in mango (*Mangifera indica* L.) Anacardiaceae with microsatellites. Mol. Breed., **15**: 383-393.

Weber J. L. 1990. Informativeness of human (dC-dA)n (dG-dT)n polymorphisms. Genomics, **7**: 524-530.

Xu X., Peng M., Fang Z. and Xu X. 2000. The direction of microsatellite mutations is dependent upon allele length. Nat. Genet., **24**: 396-399.

Yonemori K., Honsho C., Kanzaki S., Eiadthong W. and Sugiura A. 2002. Phylogenetic relationships of *Mangifera* species revealed by ITS sequences of nuclear ribosomal DNA and a possibility of their hybrid origin. Plant Syst. Evol., **231**: 59-75.